

WHAT IS CLAIMED IS:

1. A system for providing video conferencing synchronization, comprising:

an audio mixer on a first device operable to receive  
5 one or more incoming audio streams, each of the one or  
more incoming audio streams having an associated  
timestamp, the audio mixer operable to generate a mixed  
audio stream from the one or more incoming audio streams,  
the audio mixer operable to determine differences in the  
10 time base of each of the one or more incoming audio  
streams and the time base for the mixed audio stream, the  
audio mixer operable to generate mapping parameters  
associated with the determined differences, the audio  
mixer operable to transform the timestamp of each of the  
15 one or more incoming audio streams to a corresponding  
output timestamp associated with the mixed audio stream  
according to the mapping parameters;

a video mixer on a second device operable to receive  
one or more incoming video streams, each of the one or  
20 more incoming video streams having an associated  
timestamp, the video mixer operable to generate a mixed  
video stream from the one or more incoming video streams,  
the video mixer operable to receive the mapping  
parameters from the audio mixer, the video mixer operable  
25 to transform the timestamp of each of the one or more  
incoming video streams to a corresponding output  
timestamp associated with the mixed video stream  
according to the mapping parameters.

2. The system of Claim 1, further comprising:

an endpoint operable to receive the mixed audio stream from the audio mixer and the mixed video stream from the video mixer, the endpoint operable to  
5 synchronize the mixed audio stream and the mixed video stream according to their respective output timestamps.

3. The system of Claim 1, wherein the audio mixer is operable to determine when the mapping parameters are  
10 to be applied.

4. The system of Claim 1, wherein the video mixer is operable to receive speaker selection information with the mapping parameters from the audio mixer.  
15

5. The system of Claim 1, wherein the mapping parameters include an offset and a scale factor, the offset and scale factor representing differences in time bases of the one or more incoming audio streams and the  
20 mixed audio stream.

6. A method for providing video conferencing synchronization, comprising:

receiving one or more incoming audio streams, each  
of the one or more incoming audio streams being  
5 associated with endpoint local timestamps;

converting the endpoint local timestamps to endpoint  
network timestamps for each of the one or more incoming  
audio streams;

generating a mixed audio stream from the one or more  
10 incoming audio streams, the mixed audio stream having  
mixer network timestamps;

determining mapping parameters between the endpoint  
network timestamps and the mixer network timestamps.

15 7. The method of Claim 6, further comprising:

receiving information indicating a relationship  
between the endpoint local timestamps and the endpoint  
network timestamps for each of the one or more incoming  
audio streams.

20

8. The method of Claim 6, further comprising:

comparing two sets of endpoint network timestamps to  
mixer network timestamps;

identifying an offset and a scale factor as the  
25 mapping parameters in response to timestamp comparison.

9. The method of Claim 6, further comprising:

adjusting the mapping parameters in response to a  
change in average jitter buffer levels associated with  
30 each one of the incoming audio streams.

10. The method of Claim 6, further comprising:

providing the mapping parameters to video stream processing in order to synchronize the mixed audio stream to an associated mixed video stream.

5

11. A system for providing video conferencing synchronization, comprising:

means for receiving one or more incoming audio streams, each of the one or more incoming audio streams being associated with endpoint local timestamps;

10

means for converting the endpoint local timestamps to endpoint network timestamps for each of the one or more incoming audio streams;

means for generating a mixed audio stream from the one or more incoming audio streams, the mixed audio stream having mixer network timestamps;

15

means for determining mapping parameters between the endpoint network timestamps and the mixer network timestamps.

20

12. The system of Claim 11, further comprising:

means for receiving information indicating a relationship between the endpoint local timestamps and the endpoint network timestamps for each of the one or more incoming audio streams.

25

13. The system of Claim 11, further comprising:

means for comparing two sets of endpoint network timestamps to mixer network timestamps;

means for identifying an offset and a scale factor as the mapping parameters in response to timestamp comparison.

30

14. The system of Claim 11, further comprising:  
means for adjusting the mapping parameters in  
response to a change in average jitter buffer levels  
associated with each one of the incoming audio streams.

5

15. The method of Claim 11, further comprising:  
means for providing the mapping parameters to video  
stream processing in order to synchronize the mixed audio  
stream to an associated mixed video stream.

10

16. A computer readable medium including code for  
providing video conferencing synchronization, the code  
operable to:

receive one or more incoming audio streams, each of  
the one or more incoming audio streams being associated  
with endpoint local timestamps;

convert the endpoint local timestamps to endpoint  
network timestamps for each of the one or more incoming  
audio streams;

generate a mixed audio stream from the one or more  
incoming audio streams, the mixed audio stream having  
mixer network timestamps;

determine mapping parameters between the endpoint  
network timestamps and the mixer network timestamps.

25

17. The computer readable medium of Claim 16,  
wherein the code is further operable to:

receive information indicating a relationship  
between the endpoint local timestamps and the endpoint  
network timestamps for each of the one or more incoming  
audio streams.

30

18. The computer readable medium of Claim 16,  
wherein the code is further operable to:

compare two sets of endpoint network timestamps to  
mixer network timestamps;

5 identify an offset and a scale factor as the mapping  
parameters in response to timestamp comparison.

19. The computer readable medium of Claim 16,  
wherein the code is further operable to:

10 adjust the mapping parameters in response to a  
change in average jitter buffer levels associated with  
each one of the incoming audio streams.

20. The computer readable medium of Claim 16,  
15 wherein the code is further operable to:

provide the mapping parameters to video stream  
processing in order to synchronize the mixed audio stream  
to an associated mixed video stream.